STATISTICS– WORKSHEET 3

1. b) Total Variation = Residual Variation + Regression Variation
2. c) binomial
3. a) 2
4. a) Type-I error
5. d) Confidence coefficient
6. b) Increase
7. b) Hypothesis
8. d) All of the mentioned
9. a) 0
10. What Is Bayes' Theorem?

* Bayes' theorem provides us the probability of an event based on new information that is, or may be related, to that event. The formula can also be used to see how the probability of an event happening is affected by hypothetical new information, supposing the new information will turn out to be true.

In terms of mathematical equations,

P (A/B) = P (B|A) P (A)

P (B)

**Where:**

P (A) = The probability of event A

P (B) = The probability of event B

P(A∣B)=The probability of A given B

P(B∣A)= The probability of B given A

11. What is z-score?

* A Z-score is actually defined as how many standard deviations above or below the mean a data point is.

It is statistically defined as

Z= Data point- Mean/Standard deviation

A z-score is close to zero tells that the data points is close to the mean value of the dataset. A positive z-score has a value above the mean value and negative Z-score has a value that is below average of all data points.

12. What is t-test?

* A t-test is a statistical test that is used for comparing the mean of two data groups. In an investigation, a t-test used to calculate whether or not differences seen between the control and each trail group are a factor of the manipulated variable or simply the result of chance. If there is significant difference between the two groups then it is most likely due to sampling error (**Alternative hypothesis**), if there is no significant difference between the groups then your **Null Hypothesis** is true.

13. What is percentile?

* In statistics terms, a percentile is defined as a score below which a given percentage of scores in its frequency distribution fall or a score at or below which a given percentage fall For example, the 50th percentile (the median value of a dataset) is the score below which 50% or at or below which 50% of the scores in the distribution may be found.

14. What is ANOVA?

* Analysis of variance, or ANOVA, is a statistical technique that is used for showing the difference between two or more means or components through significance tests. It also shows us a way to make multiple comparisons of several populations’ means. The Anova test is performed by contrasting two types of variation, the variation between the sample means, as well as the variation within each of the samples.

15. How can ANOVA help?

* ANOVA is useful for testing three or more variables. It is close to performing multiple two-sample t-tests. Although, it outcomes in fewer type I errors and is appropriate for a range of issues. ANOVA groups differences by comparing the means of each data group and includes spreading out the variance into different origin in the statistical data set.